Source:sink imbalance detected with leaf and canopy-level spectroscopy in a field-grown crop

Authors: Angela C. Burnett, Shawn P. Serbin, Alistair Rogers

Contact Information for all authors:
Environmental and Climate Sciences Department
Brookhaven National Laboratory
Upton, New York, USA
Present address for Angela C. Burnett: Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EA, UK

Running head: Detecting source:sink imbalance using spectroscopy

Funding:
This work was supported by the United States Department of Energy with contract number DE-SC0012704 to Brookhaven National Laboratory.

Summary statement:
Reducing the carbon sink of field-grown zucchini plants alters the metabolism and structure of leaves. These changes are detectable in the light reflected from the plants, measured at both the leaf and canopy scales.

Keywords: *Cucurbita pepo*, food security, growth, development, remote sensing, metabolic profiles, carbohydrates

Acknowledgements:
This work was supported by the United States Department of Energy with contract number DE-SC0012704 to Brookhaven National Laboratory. We are grateful to Keith Lewin for preparing the experimental field and advising on agricultural practices. We acknowledge Duncan Anderson, Matthew Burnett, Sophie Drew, Casey Hamilton, Benjamin Miller and Ivanellis Rodriguez-Torres for assistance with physiological and spectral measurements, sample collection, and field management. We thank Keith Lewin and Jeremiah Anderson for designing and building the truck-mounted boom for spectral data collection. We acknowledge Jeremiah Anderson, Andrew McMahon and Daryl Yang for their assistance with UAS data collection.

Abstract

The finely-tuned balance between sources and sinks determines plant resource partitioning and regulates growth and development. Understanding and measuring metabolic indicators of source or sink limitation forms a vital part of global efforts to increase crop yield for future food security. We measured metabolic profiles of *Cucurbita pepo* (zucchini) grown in the field under carbon sink limitation and control conditions. We demonstrate that these profiles can be measured non-destructively using hyperspectral reflectance at both leaf and canopy scales. Total non-structural carbohydrates (TNC) increased 82% in sink-limited plants; leaf mass per unit area (LMA) increased 38%; free amino acids increased 22%. Partial least-squares regression models link these measured functional traits with reflectance data, enabling high-throughput estimation of traits comprising the sink limitation response. Leaf- and canopy-scale models for TNC had R^2 values of 0.93 and 0.64 and %RMSE of 13% and 38% respectively. For LMA, R^2 values were 0.91 and 0.60 and %RMSE 7% and 14%; for free amino acids, R^2 was 0.53 and 0.21 with %RMSE 20% and 26%. Remote sensing can enable accurate, rapid detection of sink limitation in the field at the leaf and canopy scale, greatly expanding our ability to understand and measure metabolic responses to stress.

Introduction

Plant breeders are faced with a significant challenge: the need to develop high-yielding crop varieties, which are resilient to future climate change (Ainsworth, Rogers & Leakey 2008b; Ort *et al.* 2015; Simkin, López-Calcagno & Raines 2019). Recent work in model crops has demonstrated that improvements in yield can be realised by engineering more efficient photosynthesis (Kromdijk *et al.* 2016; South, Cavanagh, Liu & Ort 2019; Degen, Worrall & Carmo-Silva 2020; Li *et al.* 2020; López-Calcagno *et al.* 2020). However, whilst increased CO_2 assimilation (carbon source activity) can be shown to enhance yield, the response is often not matched by an equivalent increase in plant growth and yield (carbon sink activity). Elevated CO_2 research, including free-air concentration enrichment (FACE) experiments, has shown that relatively large stimulations in photosynthesis do not always translate to commensurate increases in growth and yield providing evidence for a sink limitation bottleneck (Long, Ainsworth, Leakey, Nösberger & Ort 2006; Ainsworth, Leakey, Ort & Long 2008a; Leakey *et al.* 2009; Sanz-Sáez *et al.* 2010). Thus, carbon sink limitation can reduce the potential for enhanced yield resulting from genetically engineered improvements to source activity and the anticipated yield benefit of rising atmospheric [CO_2]. To fully capitalise on genetically enhanced and CO_2-stimulated photosynthesis, carbon sink limitation must be minimised.

Indeed, an integrated understanding of carbon sources and sinks is increasingly recognised as a vital component of enhancing global food production (White, Rogers, Rees & Osborne 2016; Smith, Rao & Merchant 2018; Fernie *et al.* 2020). If breeders are to achieve the goal of minimising sink limitation (Fernie *et al.* 2020), it is essential to understand how carbon sink limitation may be measured in a high-throughput manner, in order to facilitate rapid screening for sink limitation in breeding programs (Reynolds & Langridge 2016).

We address this need by defining the metabolic signature of carbon sink limitation in field-grown *Cucurbita pepo* and then developing a high-throughput system for measuring this sink limitation non-destructively using "hyperspectral", or high spectral resolution, reflectance data. Unlike several other limitations on growth that are of interest to breeders, such as disease resistance or drought resilience in response to pathogen or water stress, there is often no visual phenotype associated with sink limitation. This means that effective measurement of sink stress relies upon destructive harvesting, making continual monitoring difficult to achieve. Furthermore, carrying out the biochemical analysis required for detailing the sink stress profile is time consuming and expensive. For these reasons, a non-destructive, high-throughput, spectroscopic approach to monitoring sink stress is a desirable tool for crop breeders.

For the first time, we examine the source:sink balance of field-grown plants using hyperspectral data. The metabolic profile of carbon sink limitation has been extensively characterised (Stitt & Krapp 1999; Bénard *et al.* 2015; Burnett, Rogers, Rees & Osborne 2016). Chief among the key traits of a carbon sink-limited plant is an increase in leaf carbohydrate content, which has been phenomenologically and mechanistically linked to reduced sink strength (Pollock & Cairns 1991; Farrar 1996; Rogers & Ainsworth 2006; Ainsworth & Bush 2011). It has also been well established that reflectance data can be used to estimate a suite of leaf traits using a range of approaches, including empirical partial least-squares regression (PLSR) modelling to build relationships between spectral data and measured traits. This spectra-trait PLSR approach provides a rapid, high-throughput means for the estimation of biochemical traits of interest (Serbin, Singh, McNeil, Kingdon & Townsend 2014; Yendrek *et al.* 2017; Silva-Perez *et al.* 2018; Ely, Burnett, Lieberman-Cribbin, Serbin & Rogers 2019; Meacham-Hensold *et al.* 2019; Cotrozzi & Couture 2020), including those associated with source-sink balance and carbon and nitrogen status. Here, we build on these advances in two key ways: (1) we evaluate the spectroscopic approach for identifying sink limitation in the production environment and (2) we scale up from leaf- to canopy-level acquisition of hyperspectral data using a boom-mounted spectrometer – a key step in moving to truly high-throughput

monitoring using remote sensing tools. We evaluated the spectra-trait modelling approach in field-grown *C. pepo* with a sink removal treatment, to determine if the relationships between reflectance and leaf traits could be used for predicting key indicators of source:sink imbalance in the field. We also investigated the capability of linear discriminant analysis (LDA) and partial least squares discriminant analysis (PLS-DA) for identification of metabolic stress using measured leaf traits and using raw hyperspectral data, respectively. These discriminant analyses were used to build class-prediction models which demonstrated the use of trait and spectral data to distinguish between plants in the control or sink removal treatment. Importantly, we demonstrate with each of our approaches that sink limitation may be successfully detected at both the leaf and canopy scale.

We tested the following hypotheses: 1) There will be significant metabolic and structural differences between leaves of sink-limited and control field-grown *C. pepo* plants. 2) These metabolic and structural differences can be detected in the field using hyperspectral reflectance data acquired at the leaf scale. 3) These metabolic and structural differences can also be detected remotely using hyperspectral reflectance data collected at the canopy scale.

Materials and Methods

Plant material and experimental treatments

There are diverse ways to experimentally manipulate the carbon source:sink balance, including: defoliation, debudding or sink removal; manipulation of temperature, light, nitrogen or CO_2 levels; and transgenic modifications (Ainsworth, Rogers, Nelson & Long 2004; White *et al.* 2016). Direct manipulations of the carbon sink in field experiments are comparatively rare. Here, we reduced the carbon sink of field-grown *C. pepo* by removing developing fruits, throughout the duration of the experiment. The continual removal of fruits is a standard agricultural practice for harvesting zucchini, although we removed fruits early in their development.

C. pepo was selected for this experimental work because it rapidly forms a full canopy, making this species an ideal target for our proof-of-concept study of canopy reflectance. *C. pepo* is a suitable model crop for sink manipulation experiments because its fruits are highly visible due to large colorful flowers, there are relatively few fruits per plant, and the fruits are easily removed. Furthermore, the local environment is suitable for growing *C. pepo*, which is commonly cultivated as a commercial crop on Long Island.

Seeds of *C. pepo* L. var. Dunja were obtained from the Long Island Cauliflower Association (Riverhead, New York, USA) and grown in a research field at Brookhaven National Laboratory, Upton, New York, USA in 2019 (latitude 40.864466, longitude 72.875158, 18 meter elevation). Following initial pH testing, the field was prepared with lime prior to sowing to ensure an appropriate soil pH for *C. pepo.* Seeds were sown on DOY 158 at a density to achieve full canopy coverage in eighteen 10m x 10m plots, each surrounded by a border of *C. pepo*, of the same width as 1.5 times the height of mature *C. pepo* (estimated from previous experiments) to give a total sown area of 12m x 12m. The large plot size was selected to facilitate canopy- and UAS-level data collection. In addition to the sink manipulation experiment described here, the field was also used for a drought experiment. Six plots underwent sink manipulation and six plots underwent drought treatment; six plots served as controls. For the sink manipulation treatment, developing fruits were removed from each plant twice per week beginning DOY 196 by using a short-bladed serrated harvesting knife to slice the midpoint of the short fleshy stem at the base of the fruit. The use of a complete, regularly maintained sink removal treatment gave rise to large variation in metabolite contents in order to fill the 'trait space' (the numerical range of data points), facilitating the development of robust PLSR models. For the drought treatment, following germination and plant establishment irrigation was withheld from drought plots resulting in soil

dry-down during periods of no precipitation. Irrigation was maintained at standard local agricultural levels in the sink manipulation and control plots. The drought treatment began on DOY 186 and lasted until the end of the experiment. Data from drought plots were not used in the main study presented here, but were included in PLSR model to increase model performance by further increasing the range of trait values, thus giving more robust data prediction capabilities. Data were collected from each of the three plot types on every measurement date. Meteorological data are reported in Supplementary Fig. 1.

Experimental schedule

Leaf and canopy spectral data were collected twice per week for the duration of the experiment (from the initiation of the sink removal treatment until senescence of the control plants, Fig. 3A). All spectral data collection and leaf harvesting was always performed within three hours of solar noon (i.e. between 10:00 and 16:00 EDT), because canopy spectral data collection requires the sun to be high in the sky to provide even illumination of the leaves. Leaf harvests for obtaining biochemical traits were alternately paired with either leaf spectral data collection or canopy spectral data collection on any given day, in order to facilitate the development of PLSR models at both the leaf and canopy scales. Conducting full diurnal time courses of leaf

metabolite measurements was beyond the scope of this work and not possible given the constraints of sampling canopy spectra around solar noon. However, the strong experimental treatment used to build predictive models should enable the prediction of metabolite contents at different times of day via leaf level spectral data and this will be an important future application of the work presented here.

Leaf spectral data

For leaf spectral data collection, three sets of random coordinates were selected on each measurement day and applied to each plot. The newest fully-expanded leaf at each coordinate was selected for measurement. First, leaf temperature was measured using a handheld infrared radiometer (Apogee Instruments, Logan, Utah, USA). Immediately afterwards, spectral data were collected using a PSR+ full-range (continuous 350-2500nm) spectroradiometer (Spectral Evolution, Lawrence, Massachusetts, USA), connected to a leaf clip assembly with an internal, calibrated light source (SVC, Poughkeepsie, New York, USA). The spectroradiometer was calibrated using a LabSphere Spectralon® reflectance standard disc (LabSphere, Inc., North Sutton, New Hampshire, USA). For each leaf, three spectral measurements were taken across the adaxial surface, and then averaged to give a single spectrum.

Canopy spectral data

For canopy spectral data collection, three evenly spaced locations were measured within the central region of each plot. The spectroradiometer was fitted with a 14 degree lens (Spectral Evolution, Lawrence, Massachusetts, USA) and positioned 2 meters above the canopy on a truck-mounted boom (Fig. 4B). The spectroradiometer was calibrated using a LabSphere Spectralon® reflectance standard plate (LabSphere, Inc., North Sutton, New Hampshire, USA). Leaf temperature in the area viewed by the spectroradiometer was first measured using the infrared radiometer, positioned above the canopy. For each location within the plot, three to five spectral measurements were performed (immediately following temperature measurement) and then averaged to give a single spectrum. When it was not possible to measure canopy spectra on all experimental plots due to logistical or weather constraints, the omitted plots were noted and prioritised for measurement on the next measurement date, ensuring an overall even coverage of canopy spectral data for all experimental plots.

Physiological data

Measurements of Φ_{PSII} and relative chlorophyll content were performed using the PhotosynQ MultispeQ V 2.0 (Kuhlgert *et al.* 2016) on the newest fully-

218 expanded leaf immediately following leaf spectral data collection. The
219 MultispeQ measurements were made on light-adapted leaves, at ambient
220 conditions matching the incident photosynthetically active radiation (PAR)
221 and the temperature at the leaf surface, using the 'Photosynthesis RIDES'
222 protocol available online at photosynq.org.
223
224 In addition to the regular measurements of Φ_{PSII} performed on all plots
225 throughout the experiment, the response of photosynthesis to intercellular
226 CO_2 concentration (A/C_i curves) were performed on a subset of plots on four
227 dates. However, this data collection was constrained by instrument
228 availability. These data, presented in Supplementary Fig. 2, were collected in
229 the field between the hours of 06:00-14:00, using a LI-6800 Portable
230 Photosynthesis System (LI-COR Biosciences, Lincoln, Nebraska, USA). A
231 diurnal measurement of photosynthesis was performed prior to
232 measurement of A/C_i curves to ensure that measurements were completed
233 before the onset of the afternoon suppression of photosynthesis occurring
234 naturally in plants in all treatments. Light response curves were performed
235 to determine the saturating irradiance to be used in A/C_i curves: 2000 µmol
236 photons m^{-2} s^{-1}. A/C_i curves were performed on the newest fully-expanded
237 and physiologically mature leaf. Leaves were acclimated in the leaf cuvette
238 until steady-state A and g_s were reached (20-45 minutes). Each A/C_i curve
239 began at 400 µmol mol^{-1} CO_2 and the CO_2 concentration was decreased then

increased in a stepwise manner as described previously (Rogers, Serbin, Ely, Sloan & Wullschleger 2017).

Steady-state values of C_i/C_a and A_{sat} at 400 µmol mol^{-1} CO_2 were obtained from the first point of the A/C_i curve. $V_{c,max}$ was estimated from A/C_i response curves and should therefore be considered as apparent $V_{c,max}$ since mesophyll conductance was not measured, meaning that values are based on intercellular rather than chloroplastic [CO_2]. Estimation of $V_{c,max}$ was made using the kinetic parameters and their temperature dependence as presented previously (Bernacchi, Singsaas, Pimentel, Portis & Long 2001; Bernacchi *et al.* 2013) following the method described in detail by Rogers et al (2017). The average root-mean-squared error (RMSE) associated with fitting $V_{c,max}$ was 1.26 (< 1% of estimated $V_{c,max}$) ± 0.88 (standard deviation). Values of $V_{c,max}$ were normalised to 25°C using an Arrhenius function (Bernacchi *et al.* 2013).

Sample collection

Leaves were either harvested immediately following leaf spectral data collection for each plant, or following canopy spectral data collection for each area of the plot. For harvests paired with leaf spectra, the leaf that had been used for spectral data collection was harvested. For harvests paired with

canopy spectra, the newest fully-expanded leaf from the center of the area viewed by the spectroradiometer was harvested. Each leaf was divided into two equal halves along the midrib. Discs from one half were punched evenly across the leaf surface, placed into an aluminium foil packet and immediately flash frozen in liquid nitrogen, at the field site. The second half of the leaf was kept intact and sealed in a plastic bag containing a damp paper towel, to prevent desiccation during the sampling of the remaining leaves. These intact leaf halves were placed into a cooler at the field site to prevent sample deterioration. After harvesting was complete, all samples were returned to the laboratory. Frozen samples were stored at -70°C for subsequent biochemical analysis. The intact half of each leaf was punched into discs of known area distributed evenly across the leaf surface, weighed, and transferred to a drying oven for the subsequent determination of leaf mass per unit leaf area (LMA) and leaf water content (LWC).

Leaf trait analysis

Leaf mass per unit leaf area was obtained from the measured area and dry mass of oven-dried leaves. Leaf water content (LWC) was obtained from leaf fresh mass at time of harvest and leaf dry mass after oven drying, according to the following formula:

$$LWC\ (\%) = \frac{(leaf\ fresh\ mass - leaf\ dry\ mass)}{leaf\ fresh\ mass}\ x\ 100$$

Analysis of leaf carbon- and nitrogen-containing metabolites (glucose, fructose, sucrose, starch, amino acids and protein) was performed as described previously (Burnett *et al.* 2016). In brief, sequential ethanol extractions were used to extract metabolites from frozen tissue. For sugars (glucose, fructose and sucrose) a continuous enzymatic substrate assay was performed in the presence of ATP and NADP; the NADPH signal associated with each sugar was measured at 340 nm (ELx808 Plate Reader, BioTek, Winooski, VT, USA). Amino acids were quantified using fluorescamine in the presence of sodium borate buffer, with fluorescence measured at 360 nm excitation, 460 nm emission and 40 nm bandwidth (Synergy HT Plate Reader, BioTek, Winooski, VT, USA) after 5 min dark incubation. Protein was quantified from the pellets resulting from the ethanol extraction using a commercially available kit (Pierce BCA protein assay kit, Thermoscientific, Rockford, IL, USA) following solubilisation in 0.1M sodium hydroxide. For starch, pellet samples were first neutralised with hydrochloric acid following the protein assay. An overnight enzymatic digest was performed and the resultant sugars were quantified as described above. For each assay, a standard curve was included on every plate to ensure accurate metabolite quantification. For a more detailed description of these methods, refer to Burnett *et al*. (2016). Biochemical traits were expressed on a per unit area

basis, derived from the relationship between fresh mass and leaf area that was obtained from the oven-dried samples.

Values of carbohydrate-corrected LMA (ccLMA) were obtained from the total non-structural carbohydrate data (TNC; the sum of glucose, fructose, sucrose and starch) as follows. TNC, expressed as mmol glucose equivalents m^{-2}, was first multiplied by the millimolar mass of glucose to give TNC, g m^{-2}. This value was then multiplied by the area of each sample to give TNC, g $sample^{-1}$. Finally ccLMA (g m^{-2}) was obtained using the following equation:

$$ccLMA = \frac{sample\ dry\ mass - sample\ TNC\ mass}{sample\ area}$$

UAS flight data

An unoccupied aerial system (UAS) flight was performed on DOY 226, the last day of the measurement period, using an 'Osprey' system as described previously (Yang *et al.* 2020). Since the flight was carried out at the end of the measurement period, only the RGB data were analysed for this study; the full data are available online as detailed at the end of the manuscript. Green Chromatic Coordinate (GCC; Richardson 2019) was obtained for each experimental plot using RGB camera data following the methods of Yang *et*

al. (Yang *et al.* 2020). This metric enables standardisation of RGB data between different cameras, facilitating comparison with future work.

Data analysis

All data analysis was performed in the R open source software environment (R Core Team 2019). For analysis of leaf traits, depicted in Figs. 1 and 2 and Table 1, each trait was analysed using repeated-measures ANOVA. Analysis was performed at the plot level (n=6 plots for each treatment). For each measurement date, the within-plot values obtained from three leaves per plot, were first averaged to give one value for each trait per plot and per measurement date. Next, if required, data were log- or square-root-transformed prior to analysis to satisfy requirements for normally distributed data. ANOVA was used to test for effects of and interactions between treatment categories (control and sink-limited) and time into the sink manipulation experiment (DOY), with repeated measurements at the plot level. The significance level was set to $p < 0.05$, with individual significance levels reported in Table 1. A post-hoc Tukey test was performed to determine the dates upon which differences between treatments were significant.

Partial least-squares regression (PLSR) was used to predict leaf traits from spectral data using the 'pls' package (Mevik & Wehrens 2007) in R. PLSR models included drought plants, control plants and sink-limited plants to increase the predictive power by extending the range of trait values and the number of samples. Measured starch and total non-structural carbohydrate (TNC) data were square-root-transformed prior to modelling; untransformed data are always presented in the manuscript. A small set of samples was removed from the dataset prior to fitting, due to outlier residual errors, for each trait.

Leaf-level PLSR models were built using all spectral wavelengths between 500 and 2400nm with the exception of TNC and free amino acids when the range 1100-2400nm was used to improve accuracy of model prediction. Canopy-level PLSR models were built using the spectral wavelengths 500-1800nm and 1950-2400nm in order to eliminate the 1800-1950nm region containing atmospheric water interference. For canopy-level PLSR models for TNC and starch, the starting wavelength was 1100nm rather than 500nm to improve the model fit. Spectral data did not undergo any transformation prior to model building.

For all PLSR models, observational data points were subset according to treatment then randomly assigned to datasets for calibration (80% of the

data) and validation (20% of the data). Component selection and model calibration were carried out as described previously (Serbin *et al.* 2014; Ely *et al.* 2019). The R^2 and root-mean-squared error (RMSE) of prediction of the validation data set was used to assess each model and the variable importance of projection (VIP) was used for qualitative evaluation of model predictor variables as described previously (Wold, Sjöström & Eriksson 2001).

Partial least squares discriminant analysis (PLS-DA) was performed in R using the 'caret' package (Kuhn 2008), in accordance with methods developed previously (Serbin *et al.* 2014; Ely *et al.* 2019; Cotrozzi & Couture 2020; Gold *et al.* 2020). PLS-DA models used 75% of the data for model training and 25% for model testing, with 10-fold cross-validated resampling repeated five times, and receiver-operator curves (ROC)-optimisation of the number of components. Models ran until convergence was reached (up to 100 iterations). The linear discriminant analysis (LDA) model was built using 75% of the data for model training and 25% for model testing; the LDA model was trained with leave-one-out calibration and ROC-optimisation of the number of components. The average results from ten model iterations are reported.

GCC was analysed using a t-test of plot-level mean GCC values for sink-limited and control plots (Supplementary Fig. 3).

Results

Photosynthesis is maintained in sink-limited plants

The sink removal treatment began on day of year (DOY) 196 and plants were measured from the onset of treatment until DOY 226 when control plants were senescing. Physiological, metabolic and structural traits were measured (Figs. 1, 2). Averaged over the experiment, leaf temperature was 3% higher in sink limited plants; leaf temperature was 6% higher in sink limited plants at the final time point. Temperature differences were significant across the experiment taken as a whole; when individual measurement dates were analysed, significant differences in temperature occurred on DOY 205, 206, 214, 221 (Fig. 1A; Table 1). Photosystem II operating efficiency (Fig. 1B; Table 1) was not affected by the sink manipulation. Leaf chlorophyll content was not affected by the sink manipulation except for DOY 200 when it was significantly higher in sink manipulated plants (Fig. 1C; Table 1). The lack of overall photosynthetic response displayed in data collected throughout the experiment (Fig. 1B) was supported by a small dataset of A/C_i curves which showed no clear trend in A_{sat} or $V_{c,max}$ although g_s was lower in sink-limited plants ($F_{1,8} = 7.2$, $p < 0.05$; Supplementary Fig. 2).

Sink limitation affects leaf metabolism and structure

Leaf structure was markedly affected by the sink manipulation treatment. There was a 5% decrease in leaf water content (LWC) in the sink manipulation treatment and the magnitude of this effect increased with time, with a highly significant time x treatment interaction (Fig. 2A; Table 1). We observed a 38% increase in LMA (mean for all time points) overall, displaying the highest increase of 57% at the final time point, and a highly significant time x treatment interaction (Fig. 2B; Table 1). The difference between control and sink-limited plants was significant on DOY 200 for LMA, and for both LMA and LWC on all measurement dates from DOY 205 onwards. About half of the overall increase in the raw LMA data presented here was attributable to an increased total non-structural carbohydrate (TNC) content in sink-limited plants; carbohydrate-corrected values of LMA (ccLMA) still showed an overall 20% increase in sink-limited plants (see Supplementary Data).

All leaf metabolites were analysed on an area basis. Free amino acid content increased by 22% in sink-limited plants and decreased over time in both control and sink-limited plants with no significant interaction (Fig. 2C; Table 1). However, for individual dates the difference was significant on DOY 200, 205, 207, 214 and 218.

Sink-limited plants displayed a strong and highly significant increase in TNC, which became more marked over time (Fig. 2H; Table 1). The increase in TNC was attributable in part to an increased sugar content (Fig. 2D,E,F; Table 1) but was dominated by a marked increase in starch (Fig. 2G; Table 1). Glucose was significantly higher in sink-limited plants on DOY 207, 214, 218, 221; fructose was significantly lower in sink-limited plants on DOY 200 then significantly higher on DOY 214, 218, 221; sucrose was significantly higher in sink-limited plants on DOY 211, 214, 218 and 221. Both starch and TNC were significantly higher in sink-limited plants on all measurement dates from DOY 205 onwards. Overall, TNC increased by 82% in sink-limited plants. There was a highly significant time x treatment interaction for each carbohydrate measured; the difference between control and sink-limited plants increased as time progressed (Fig. 2D,E,F,G,H; Table 1).

Leaf protein content increased 8% overall in sink-limited plants, and the magnitude of this difference was greatest at the final time point when protein was 25% higher than in the control plants; there was a significant time x treatment interaction and the difference between treatments was significant on DOY 207, 218 and 221 (Fig. 2I; Table 1).

Plot-level greenness is maintained in sink-limited plants

Sink-limited plots stayed green for longer than control plots at the end of the experiment, due to delayed senescence (Fig. 3). Using UAS imagery, we observed that the Green Chromatic Coordinate (GCC) was significantly higher in sink-limited than control plants ($t = 5.8$, $p < 0.001$, $df = 10$; Supplementary Fig. 3). This response was also visually evident in the standard red-green-blue (RGB) image where sink-limited plots were visually greener compared to the control plots (Fig. 3A).

Partial least-squares regression successfully predicts metabolic and structural traits from reflectance at leaf and canopy scales

Reflectance data were collected at both leaf and canopy scales using a leaf clip (Fig. 4A) and truck-mounted boom (Fig. 4B). At both scales, partial least-squares regression (PLSR) successfully estimated leaf metabolite contents and structural traits associated with sink limitation (Fig. 5). In general, models performed better at the leaf level, with a higher R^2 when compared to the canopy-level model for the same trait, and a lower RMSE for six of nine leaf-level models when compared to canopy-level models (Fig. 5; data shown are for independent model validation in each case). R^2 values for leaf-level models ranged from 0.53 to 0.93 demonstrating strong

predictive capabilities (Fig. 5). Canopy-level models also showed acceptable predictive capabilities for many traits. Five of the nine models had R^2 values greater than 0.5 (Fig. 5) with the highest R^2 of 0.78 for LWC. The %RMSE (RMSE expressed as a percentage of the mean of the observed values for a trait) ranged from 2% to 37% for leaf models, and from 2% to 38% for canopy models. For leaf models, %RMSE was <20% for all traits except sugars, and was 2% for LWC, 7% for LMA, 20% for free amino acids, 11% for protein and 13% for TNC. For canopy models, %RMSE was 2% for LWC, 14% for LMA, 26% for free amino acids, 10% for protein and 38% for TNC.

Sink stress detection may be achieved using measured traits or hyperspectral reflectance

Linear discriminant analysis (LDA) was used to determine whether or not plants were exposed to sink stress, based on metabolic and structural traits (glucose, fructose, sucrose, starch, protein, free amino acids, LMA and LWC). When LDA was performed iteratively, including cumulative data for each successive date and the preceding dates, class detection accuracy improved over time as the treatment effect became stronger. The maximum overall accuracy was 86% when all time points were included (Fig. 6), and the area under the receiver-operator curve (AUC-ROC) was 0.93. Partial least squares discriminant analysis (PLS-DA) using raw spectral data also

showed a good capability for distinguishing between sink-limited and control plants, with detection success of 78% at the leaf level (AUC-ROC = 0.86) and 89% for canopy level spectra (AUC-ROC = 0.96), including all measured time points (Fig. 6). The greater success of detection with canopy spectra is likely due to the fact that compared to leaf-scale data collection, canopy spectral data collection began slightly later into the experiment. This would enhance the overall treatment effect observed in canopy data, since the metabolic differences between treatments generally increased over time as the sink stress became more pronounced. After we omitted the leaf measurements that did not overlap with those from the canopy collections in the PLS-DA (i.e. leaf and canopy measurement periods were aligned, with the earliest part of the experiment omitted) the detection accuracy at the leaf scale was 93% (AUC-ROC = 0.99). The equivalent measurement for LDA using measured leaf traits yielded a prediction accuracy of 94% (Fig. 6). Measured traits and hyperspectral reflectance are both successful at distinguishing between sink-limited and control plants (Fig. 6).

Discussion

We conducted a sink manipulation experiment in field-grown *C. pepo* and demonstrated that we could detect the marked and significant effect of sink limitation (hypothesis 1) on leaf metabolic and structural traits using spectroscopy. Our key finding was that this approach can be scaled effectively from the leaf level to the canopy scale (hypotheses 2 and 3). Collectively our results demonstrate the robustness of the spectroscopy approach, the potential to detect sink limitation non-destructively and remotely, and to do that in a real-world agricultural setting emphasising the value of the approach for breeders and producers.

The metabolic signature of sink stress

Sink strength is the product of sink size multiplied by sink activity (Geiger & Shieh 1993; White *et al.* 2016). Removing developing fruits dramatically decreases carbon sink strength within the plant, by removing a critical carbon sink. However, fruit removal also increases the sink activity by stimulating the development of new fruits, which have a strong carbon requirement, thereby increasing sink strength within the plant. In the manipulation performed here, the net effect on sink strength was an overall decrease, since the increase in sink activity was outweighed by the larger

539 decrease in sink size, and this is confirmed by the trends in carbohydrate
540 levels observed in sink-limited plants (Fig. 2D,E,F,G,H).
541
542 Like most biotic and abiotic plant stresses, sink limitation has a metabolic
543 signature. Here, sink limitation led to significant increases in the content of
544 non-structural carbohydrates (Fig. 2D,E,F,G,H) and free amino acids (Fig.
545 2C) in addition to leaf structural changes (Fig. 2A,B). Increased levels of
546 non-structural carbohydrates – both each carbohydrate metabolite
547 individually, and the total pool (TNC) – is consistent with the literature on
548 carbon sink limitation; our sink removal treatment for manipulation of the
549 source:sink balance therefore elicited the expected response. Sink-limited
550 plants had increased levels of leaf carbohydrates likely due to decreased
551 export from the leaf caused by reduced sink demand (Stitt & Krapp 1999;
552 Ainsworth & Bush 2011; Burnett *et al.* 2016).
553
554 The observed leaf structural changes (LWC and LMA; Fig. 2A,B; Table 1) are
555 consistent with the development of smaller, longer-lasting leaves;
556 furthermore, these leaf characteristics are themselves consistent with the
557 delayed leaf senescence and maintained leaf protein content observed in
558 sink-limited plants (Figs. 2I and 3). Delayed leaf senescence has been
559 observed in multiple FACE experiments in which elevated [CO_2] increased
560 the carbon source:sink balance nondestructively, further indicating that our

findings are commensurate with a source:sink imbalance (Kontunen-Soppela *et al.* 2010; McGrath, Karnosky & Ainsworth 2010; Tallis *et al.* 2010).

Detection of metabolic traits at leaf and canopy scales using hyperspectral reflectance

Since the sink limitation treatment was not accompanied by a change in photosynthesis (Figs. 1,2; Table 1; Supplementary Fig. 2), screening for sink limitation in zucchini plants requires insight into the metabolic response. Biochemical measurements are not only time-consuming and costly to perform but are also destructive, meaning that a leaf-level study cannot track an individual leaf over its lifespan. In addition, the delay in obtaining results from destructive analysis is substantial preventing rapid feedback to breeders or farmers. Therefore the use of high-throughput, non-destructive hyperspectral data, which can readily be analysed to understand the metabolic status of a plant (at the leaf level) or field plot (at the canopy level), enables a great step forward in the efficiency and capability of sink stress monitoring. Our study provides the first field-level example of non-destructive monitoring of sink stress via metabolite prediction.

Whilst leaf-level PLSR models generally had a higher capability for predicting traits from spectral data than their canopy-level counterparts (Fig. 5),

583 models at both scales were effective at predicting a suite of leaf traits.
584 Importantly, predictions of starch, which forms the vast majority of TNC,
585 and leaf structural traits (LWC and LMA) were successful at both leaf and
586 canopy levels with R^2 > 0.60 in each case (Fig. 5A,B,G,H). Since TNC was
587 the major metabolic indicator of sink limitation in this study, this indicates
588 effective prediction of sink limitation at the canopy scale. To our knowledge
589 this is the first time that canopy-level predictive models have been used to
590 examine the traits underpinning crop sink limitation.
591
592 **Perspectives on scaling up trait detection**
593
594 Scaling detection of traits from the leaf level to the canopy level is a critical
595 step to enable high throughput measurement of plant traits (Asner & Martin
596 2008; Kokaly, Asner, Ollinger, Martin & Wessman 2009; Virlet,
597 Sabermanesh, Sadeghi-Tehran & Hawkesford 2017; Herrmann *et al.* 2018)
598 To measure traits at the canopy level, the time of day must be carefully
599 considered given the reliance on natural illumination of the leaves by solar
600 irradiation, rather than artificial illumination from the light sources typically
601 used in a leaf clip. Leaf orientation and canopy structure also become
602 relevant for canopy-scale measurements of reflectance (Ollinger 2011); gaps
603 between plants must be avoided in order to obtain a reliable spectral
604 measurement of the core vegetation component. Finally, atmospheric water

605 vapor can interfere with the signal in the main water absorption regions
606 (Gao, Heidebrecht & Goetz 1993) and must be removed from the spectral
607 data prior to analysis. Here, we successfully demonstrated the use of
608 canopy-level spectral data for detecting sink limitation, representing a major
609 advance in the phenotyping of sink limitation – a recognised critical target
610 for crop breeding (Dusenge, Duarte & Way 2019; Fernie *et al.* 2020).
611
612 Scaling the hyperspectral monitoring of sink stress to the UAS level using
613 UAS-mounted hyperspectral sensors (Yang *et al.* 2017, 2020; Shiklomanov
614 *et al.* 2019) is the next step for increasing throughput and the ability to
615 scale the technique, as has been shown for phenotyping of wheat height in
616 response to a nitrogen treatment (Holman *et al.* 2016) and canopy
617 characteristics of avocado trees (Tu, Johansen, Phinn & Robson 2019).
618 However, it must be noted that scaling up to the UAS level is not without its
619 technical, economic, and legislative challenges (Hunt & Daughtry 2018;
620 Coops, Goodbody & Cao 2019). In terms of technical limitations, UAS-
621 mounted spectral cameras often have a narrower range of wavebands and a
622 lower waveband resolution than hyperspectral sensors used on the ground,
623 e.g. 10 nm resolution in the study by Basso, Fiorentino, Cammarano &
624 Schulthess (2016), reducing measurement precision. In the present study,
625 we used the Green Chromatic Coordinate (GCC), which is a simple metric
626 derived from RGB camera data, to demonstrate that sink-limited plots were

627 greener than control plots at the end of the experimental period, measured
628 at the UAS level. This finding is likely related to underlying physiological and
629 biochemical traits, since GCC is linked to pigments and plant health as well
630 as leaf area index (Liu *et al*. 2015; Reid *et al*. 2016; Liu *et al*. 2018). Simple
631 metrics such as GCC provide a less expensive approach to airborne crop
632 monitoring, and UAS measurements frequently rely on spectral indices
633 rather than taking the full-spectrum trait prediction approach demonstrated
634 using our leaf- and canopy-level data. However, for detecting small changes
635 and understanding the underlying metabolic differences, hyperspectral data
636 provide far more detailed information than spectral indices. Hyperspectral
637 data – in contrast to multispectral data – are especially suited to
638 measurements of nutrient status as well as other stresses such as pathogens
639 when using UAS systems (Maes & Steppe 2019) and it will be important to
640 use high spectral resolution when scaling up our detailed approach for stress
641 detection from the leaf- and canopy-level to the UAS level.
642
643 The detection of metabolic and structural signatures using leaf reflectance
644 facilitates faster screening for crop breeding as well as the development of
645 precision agriculture techniques. In the plant breeding context,
646 understanding sink limitation – whether at the leaf, canopy or field scale –
647 enables the development of crops better able to translate additional
648 photosynthate resulting from improved carbon assimilation (Kromdijk *et al.*

649 2016; South *et al.* 2019; Degen *et al.* 2020; Li *et al.* 2020; López-Calcagno
650 *et al.* 2020) or future elevated CO_2 (Ainsworth *et al.* 2008b; Leakey *et al.*
651 2009) into enhanced yield. In the precision agriculture context, monitoring
652 sink limitation in major crops may be used to inform the timing of fertiliser
653 application to improve the balance between carbon and nitrogen resources in
654 the plant (Basso *et al.* 2016; Maresma, Ariza, Martínez, Lloveras & Martínez-
655 Casasnovas 2016; Maes & Steppe 2019). Both carbon and nitrogen can
656 place limits on crop growth, development and yield (Burnett *et al.* 2016;
657 Burnett, Rogers, Rees & Osborne 2018; White *et al.* 2016), and both source
658 and sink limitations must be addressed for successful breeding of our future
659 crops (White *et al.* 2016; Fernie *et al.* 2020).
660
661 In summary, source:sink balance underpins plant growth and survival and is
662 a key factor affecting crop yield. In order to realise crop yield increases, an
663 integrated understanding of carbon and nitrogen sources and sinks is
664 essential. Remote sensing provides a unique opportunity for detailed, high-
665 throughput phenotyping of plant physiological and metabolic traits, enabling
666 us to understand limitations on yield caused by sink limitation. Here we have
667 demonstrated the use of leaf reflectance data to examine vital plant
668 processes in the production environment, measuring source:sink balance
669 remotely in field-grown plants for the first time, enabling rapid and non-
670 invasive measurements of sink limitation.

Acknowledgements

This work was supported by the United States Department of Energy with contract number DE-SC0012704 to Brookhaven National Laboratory. We are grateful to Keith Lewin for preparing the experimental field and advising on agricultural practices. We acknowledge Duncan Anderson, Matthew Burnett, Sophie Drew, Casey Hamilton, Benjamin Miller and Ivanellis Rodriguez-Torres for assistance with physiological and spectral measurements, sample collection, and field management. We thank Keith Lewin and Jeremiah Anderson for designing and building the truck-mounted boom for spectral data collection. We acknowledge Jeremiah Anderson, Andrew McMahon and Daryl Yang for their assistance with UAS data collection.

Supplemental Material

Supplementary Figure 1. Meteorological data.

Supplementary Figure 2. Gas exchange data.

Supplementary Figure 3. Green chromatic coordinate data.

Supplementary Datasets. The spectra and trait data presented in this manuscript are available on-line [http://ecosis.org] from the Ecological Spectral Information System (EcoSIS) at https://doi.org/10.21232/RLmYbmE3

The UAS dataset used in this manuscript is available on-line at https://doi.org/10.17605/OSF.IO/SN6EM

Additionally, the spectra and trait data presented here are available as supplementary files accompanying this manuscript, along with the carbohydrate-corrected LMA dataset, metadata, and a ReadMe text file.

Conflict of Interest Statement

The authors have no conflict of interest to declare.

Data Availability Statement

The data used in this manuscript are publicly available free of charge in online repositories, as detailed in the Supplemental Material section.

References

Ainsworth E.A. & Bush D.R. (2011) Carbohydrate export from the leaf: a highly regulated process and target to enhance photosynthesis and productivity. *Plant Physiology* **155**, 64–9.

Ainsworth E.A., Leakey A.D.B., Ort D.R. & Long S.P. (2008a) FACE-ing the facts: inconsistencies and interdependence among field, chamber and modeling studies of elevated [CO_2] impacts on crop yield and food supply. *New Phytologist* **179**, 5–9.

Ainsworth E.A., Rogers A. & Leakey A.D.B. (2008b) Targets for crop biotechnology in a future high-CO_2 and high-O_3 world. *Plant Physiology* **147**, 13–9.

Ainsworth E.A., Rogers A., Nelson R. & Long S.P. (2004) Testing the "source–sink" hypothesis of down-regulation of photosynthesis in elevated [CO_2] in the field with single gene substitutions in *Glycine max.* *Agricultural and Forest Meteorology* **122**, 85–94.

Asner G.P. & Martin R.E. (2008) Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment* **112**, 3958–3970.

Basso B., Fiorentino C., Cammarano D. & Schulthess U. (2016) Variable rate nitrogen fertilizer response in wheat using remote sensing. *Precision Agriculture* **17**, 168–182.

Bénard C., Bernillon S., Biais B., Osorio S., Maucourt M., Ballias P., … Moing A. (2015) Metabolomic profiling in tomato reveals diel compositional changes in fruit affected by source-sink relationships. *Journal of Experimental Botany* **66**, 3391–3404.

Bernacchi C.J., Bagley J.E., Serbin S.P., Ruiz-Vera U.M., Rosenthal D.M. & Vanloocke A. (2013) Modelling C_3 photosynthesis from the chloroplast to the ecosystem. *Plant, Cell & Environment* **36**, 1641–57.

Bernacchi C.J., Singsaas E.L., Pimentel C., Portis A.R. & Long S.P. (2001) Improved temperature response functions for models of Rubisco-limited photosynthesis. *Plant, Cell and Environment* **24**, 253–259.

Burnett A.C., Rogers A., Rees M. & Osborne C.P. (2016) Carbon source–sink limitations differ between two species with contrasting growth strategies. *Plant, Cell and Environment* **39**, 2460–2472.

Burnett A.C., Rogers A., Rees M. & Osborne C.P. (2018) Nutrient sink limitation constrains growth in two barley species with contrasting growth strategies. *Plant Direct* **2**.

Coops N.C., Goodbody T.R.H. & Cao L. (2019) Four steps to extend drone use in research. *Nature* **572**, 433–435.

Cotrozzi L. & Couture J.J. (2020) Hyperspectral assessment of plant responses to multi-stress environments: Prospects for managing protected agrosystems. *Plants, People, Planet* **2**, 244–258.

Degen G.E., Worrall D. & Carmo-Silva E. (2020) An isoleucine residue acts as a thermal and regulatory switch in wheat Rubisco activase. *Plant Journal* **103**, 742–751.
Dusenge M.E., Duarte A.G. & Way D.A. (2019) Plant carbon metabolism and climate change: elevated CO_2 and temperature impacts on photosynthesis, photorespiration and respiration. *New Phytologist* **221**, 32–49.
Ely K.S., Burnett A.C., Lieberman-Cribbin W., Serbin S.P. & Rogers A. (2019) Spectroscopy can predict key leaf traits associated with source-sink balance and carbon-nitrogen status. *Journal of Experimental Botany* **70**, 1789–1799.
Farrar J.F. (1996) Sinks--integral parts of a whole plant. *Journal of Experimental Botany* **47 Spec No**, 1273–9.
Fernie A.R., Bachem C.W.B., Helariutta Y., Neuhaus H.E., Prat S., Ruan Y.L., … Sonnewald U. (2020) Synchronization of developmental, molecular and metabolic aspects of source–sink interactions. *Nature Plants* **6**, 55–66.
Gao B.C., Heidebrecht K.B. & Goetz A.F.H. (1993) Derivation of scaled surface reflectances from AVIRIS data. *Remote Sensing of Environment* **44**, 165–178.
Geiger D.R. & Shieh W. (1993) Sink strength: learning to measure, measuring to learn. *Plant, Cell & Environment* **16**, 1017–1018.
Gold K.M., Townsend P.A., Chlus A., Herrmann I., Couture J.J., Larson E.R. & Gevens A.J. (2020) Hyperspectral measurements enable pre-symptomatic detection and differentiation of contrasting physiological effects of late blight and early blight in potato. *Remote Sensing* **12**.
Herrmann I., Vosberg S.K., Ravindran P., Singh A., Chang H.X., Chilvers M.I., … Townsend P.A. (2018) Leaf and canopy level detection of Fusarium virguliforme (sudden death syndrome) in soybean. *Remote Sensing* **10**, 1–19.
Holman F.H., Riche A.B., Michalski A., Castle M., Wooster M.J. & Hawkesford M.J. (2016) High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. *Remote Sensing* **8**.
Hunt E.R. & Daughtry C.S.T. (2018) What good are unmanned aircraft systems for agricultural remote sensing and precision agriculture? *International Journal of Remote Sensing* **39**, 5345–5376.
Kokaly R.F., Asner G.P., Ollinger S. V., Martin M.E. & Wessman C.A. (2009) Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment* **113**, S78–S91.
Kontunen-Soppela S., Parviainen J., Ruhanen H., Brosché M., Keinänen M., Thakur R.C., … Vapaavuori E. (2010) Gene expression responses of paper birch (*Betula papyrifera*) to elevated CO_2 and O_3 during leaf

maturation and senescence. *Environmental Pollution* **158**, 959–968.
Kromdijk J., Głowacka K., Leonelli L., Gabilly S.T., Iwai M., Niyogi K.K. & Long S.P. (2016) Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science* **354**, 857–861.
Kuhlgert S., Austic G., Zegarac R., Osei-Bonsu I., Hoh D., Chilvers M., … Kramer D. (2016) MultispeQ Beta: a tool for large-scale plant phenotyping connected to the open PhotosynQ network. *Royal Society Open Science* **3**, 160592.
Kuhn M. (2008) Building predictive models in R using the caret package. *Journal of Statistical Software* **28**, 1–26.
Leakey A.D.B., Ainsworth E.A., Bernacchi C.J., Rogers A., Long S.P. & Ort D.R. (2009) Elevated CO_2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *Journal of Experimental Botany* **60**, 2859–76.
Li X., Wang P., Li J., Wei S., Yan Y., Yang J., … Zhou W. (2020) Maize GOLDEN2-LIKE genes enhance biomass and grain yields in rice by improving photosynthesis and reducing photoinhibition. *Communications Biology* **3**, 151.
Liu Z., An S., Lu X, Hu H. & Tang J. (2018) Using canopy greenness index to identify leaf ecophysiological traits during the foliar senescence in an oak forest. *Ecosphere* **9**, e02337.
Liu Z., Hu H., Yu H., Yang X., Yang H., Ruan C., Wang Y. & Tang J. (2015) Relationship between leaf physiologic traits and canopy color indices during the leaf expansion period in an oak forest. *Ecosphere* **6**, 1-9.
Long S.P., Ainsworth E.A., Leakey A.D.B., Nösberger J. & Ort D.R. (2006) Food for thought: lower-than-expected crop yield stimulation with rising CO_2 concentrations. *Science* **312**, 1918–21.
López-Calcagno P.E., Brown K.L., Simkin A.J., Fisk S.J., Vialet-Chabrand S., Lawson T. & Raines C.A. (2020) Stimulating photosynthetic processes increases productivity and water-use efficiency in the field. *Nature Plants* **6**, 1054–1063.
Maes W.H. & Steppe K. (2019) Perspectives for remote sensing with Unmanned Aerial Vehicles in precision agriculture. *Trends in Plant Science* **24**, 152–164.
Maresma Á., Ariza M., Martínez E., Lloveras J. & Martínez-Casasnovas J.A. (2016) Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAV service. *Remote Sensing* **8**.
McGrath J.M., Karnosky D.F. & Ainsworth E.A. (2010) Spring leaf flush in aspen (*Populus tremuloides*) clones is altered by long-term growth at elevated carbon dioxide and elevated ozone concentration. *Environmental Pollution* **158**, 1023–1028.
Meacham-Hensold K., Montes C.M., Wu J., Guan K., Fu P., Ainsworth E.A., … Bernacchi C.J. (2019) High-throughput field phenotyping using

hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. *Remote Sensing of Environment* **231**, 111176.
Mevik B. & Wehrens R. (2007) The pls Package: Principal Component and Partial Least Squares Regression in R. *Journal of Statistical Software* **18**, 1–23.
Ollinger S. V. (2011) Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist* **189**, 375–394.
Ort D.R., Merchant S.S., Alric J., Barkan A., Blankenship R.E., Bock R., … Zhu X.G. (2015) Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proceedings of the National Academy of Sciences* **112**, 8529–8536.
Pollock C.J. & Cairns A.J. (1991) Fructan metabolism in grasses and cereals. *Annual Review of Plant Physiology and Plant Molecular Biology* **42**, 77–101.
R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Reid A.M., Chapman W.K., Prescott C.E. & Nijland W. (2016) Using excess greenness and green chromatic coordinate colour indices from aerial images to assess lodgepole pine vigour, mortality and disease occurrence. *Forest Ecology and Management* **374**, 146-153.
Reynolds M. & Langridge P. (2016) Physiological breeding. *Current Opinion in Plant Biology* **31**, 162–171.
Richardson, A.D. (2019) Tracking seasonal rhythms of plants in diverse ecosystems with digital camera imagery. *New Phytologist* **222**, 1742-1750.
Rogers A. & Ainsworth E.A. (2006) The response of foliar carbohydrates to elevated [CO_2]. *Managed Ecosystems and CO_2* **187**, 293–308.
Rogers A., Serbin S.P., Ely K.S., Sloan V.L. & Wullschleger S.D. (2017) Terrestrial biosphere models underestimate photosynthetic capacity and CO_2 assimilation in the Arctic. *New Phytologist* **216**, 1090–1103.
Sanz-Sáez Á., Erice G., Aranjuelo I., Nogués S., Irigoyen J.J. & Sánchez-Díaz M. (2010) Photosynthetic down-regulation under elevated CO_2 exposure can be prevented by nitrogen supply in nodulated alfalfa. *Journal of Plant Physiology* **167**, 1558–1565.
Serbin S.P., Singh A., McNeil B.E., Kingdon C.C. & Townsend P.A. (2014) Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecological Applications* **24**, 1651–1669.
Shiklomanov A.N., Bradley B.A., Dahlin K.M., M Fox A., Gough C.M., Hoffman F.M., … Smith W.K. (2019) Enhancing global change experiments through integration of remote-sensing techniques. *Frontiers in Ecology and the Environment* **17**, 215–224.
Silva-Perez V., Molero G., Serbin S.P., Condon A.G., Reynolds M.P., Furbank

R.T. & Evans J.R. (2018) Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *Journal of Experimental Botany* **69**, 483–496.
Simkin A.J., López-Calcagno P.E. & Raines C.A. (2019) Feeding the world: Improving photosynthetic efficiency for sustainable crop production. *Journal of Experimental Botany* **70**, 1119–1140.
Smith M.R., Rao I.M. & Merchant A. (2018) Source-sink relationships in crop plants and their influence on yield development and nutritional quality. *Frontiers in Plant Science* **871**, 1–10.
South P.F., Cavanagh A.P., Liu H.W. & Ort D.R. (2019) Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science* **363**.
Stitt M. & Krapp A. (1999) The interaction between elevated carbon dioxide and nitrogen nutrition: the physiological and molecular background. *Plant, Cell & Environment* **22**, 583–621.
Tallis M.J., Lin Y., Rogers A., Zhang J., Street N.R., Miglietta F., … Taylor G. (2010) The transcriptome of *Populus* in elevated CO_2 reveals increased anthocyanin biosynthesis during delayed autumnal senescence. *New Phytologist* **186**, 415–428.
Tu Y.H., Johansen K., Phinn S. & Robson A. (2019) Measuring canopy structure and condition using multi-spectral UAS imagery in a horticultural environment. *Remote Sensing* **11**, 15–17.
Virlet N., Sabermanesh K., Sadeghi-Tehran P. & Hawkesford M.J. (2017) Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. *Functional Plant Biology* **44**, 143–153.
White A.C., Rogers A., Rees M. & Osborne C.P. (2016) How can we make plants grow faster? A source–sink perspective on growth rate. *Journal of Experimental Botany* **67**, 31–45.
Wold S., Sjöström M. & Eriksson L. (2001) PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **58**, 109–130.
Yang D., Meng R., Morrison B.D., Mcmahon A., Hantson W., Hayes D., … Serbin S. (2020) A multi-sensor Unoccupied Aerial System improves characterization of vegetation composition and canopy properties in the Arctic tundra. *Remote Sensing* **12**, 2638.
Yang G., Liu J., Zhao C., Li Z., Huang Y., Yu H., … Yang H. (2017) Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives. *Frontiers in Plant Science* **8**.
Yendrek C.R., Tomaz T., Montes C.M., Cao Y., Morse A.M., Brown P.J., … Ainsworth E.A. (2017) High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance. *Plant Physiology* **173**, 614–626.

Tables

Table 1 Leaf traits shown in Figs. 1 and 2 analysed with repeated-measures ANOVA. Effects of treatment (control and sink manipulation), time (DOY) and the interactive effect are shown.

Trait	Effect	F-value	P-value
Leaf temperature	Treatment	$F_{(1,8)} = 20.1$	$P < 0.01$
	Time	$F_{(10,94)} = 59.6$	$P < 0.001$
	Time x Treatment	$F_{(10,94)} = 1.1$	ns
Φ_{PSII}	Treatment	$F_{(1,10)} = 0.3$	ns
	Time	$F_{(6,60)} = 29.2$	$P < 0.001$
	Time x Treatment	$F_{(6,60)} = 1.4$	ns
Chlorophyll	Treatment	$F_{(1,10)} = 1.1$	ns
	Time	$F_{(6,60)} = 32.5$	$P < 0.001$
	Time x Treatment	$F_{(6,60)} = 1.5$	ns
Leaf water content	Treatment	$F_{(1,6)} = 104.5$	$P < 0.001$
	Time	$F_{(8,63)} = 27.4$	$P < 0.001$
	Time x Treatment	$F_{(8,63)} = 6.0$	$P < 0.001$
LMA	Treatment	$F_{(1,6)} = 190.7$	$P < 0.001$
	Time	$F_{(8,63)} = 53.3$	$P < 0.001$
	Time x Treatment	$F_{(8,63)} = 21.8$	$P < 0.001$
Free amino acids	Treatment	$F_{(1,8)} = 14.8$	$P < 0.01$
	Time	$F_{(7,61)} = 4.9$	$P < 0.001$
	Time x Treatment	$F_{(7,61)} = 1.4$	ns
Glucose	Treatment	$F_{(1,8)} = 13.0$	$P < 0.01$
	Time	$F_{(7,61)} = 13.6$	$P < 0.001$
	Time x Treatment	$F_{(7,61)} = 12.3$	$P < 0.001$
Fructose	Treatment	$F_{(1,8)} = 1.4$	ns
	Time	$F_{(7,61)} = 20.1$	$P < 0.001$
	Time x Treatment	$F_{(7,61)} = 9.1$	$P < 0.001$
Sucrose	Treatment	$F_{(1,8)} = 12.4$	$P < 0.01$
	Time	$F_{(7,61)} = 14.3$	$P < 0.001$
	Time x Treatment	$F_{(7,61)} = 3.8$	$P < 0.01$
Starch	Treatment	$F_{(1,8)} = 134.9$	$P < 0.001$
	Time	$F_{(7,61)} = 35.5$	$P < 0.001$
	Time x Treatment	$F_{(7,61)} = 20.4$	$P < 0.001$
TNC	Treatment	$F_{(1,8)} = 199.8$	$P < 0.001$
	Time	$F_{(7,61)} = 37.4$	$P < 0.001$
	Time x Treatment	$F_{(7,61)} = 28.8$	$P < 0.001$
Protein	Treatment	$F_{(1,8)} = 10.9$	$P < 0.05$
	Time	$F_{(7,61)} = 29.8$	$P < 0.001$
	Time x Treatment	$F_{(7,61)} = 2.3$	$P < 0.05$

Table 2 Numbers of datapoints in calibration (cal.) and validation (val.) datasets and number of model components (nComps) for partial least square regression (PLSR) models presented in Fig. 5.

	Leaf-scale PLSR			Canopy-scale PLSR		
Trait	**Cal.**	**Val.**	**nComps**	**Cal.**	**Val.**	**nComps**
LWC	192	49	7	71	20	9
LMA	190	48	9	71	19	11
Free amino acids	187	48	10	70	19	10
Glucose	191	49	11	70	18	6
Fructose	189	48	11	68	18	5
Sucrose	188	49	9	70	18	9
Starch	180	47	10	67	19	6
TNC	178	47	9	67	19	6
Protein	189	50	8	71	19	8

Figure Legends

Figure 1 Leaf temperature was higher in sink-limited plants, but photosynthesis-related traits did not change. A, leaf temperature; B, efficiency of photosystem II (ΦPSII); C, relative chlorophyll content. For leaf temperature (A), n = 6 plots each reporting an average value of 3-6 reps per plot measured at leaf and/or canopy scales on any given date; the final two time points show the average of 5 sink-limited plots and 4 control plots. For MultispeQ measurements (B,C), n = 6 plots with an average of 3 reps taken for the plot-level value. Means ± standard error of plot-level data are shown.

Figure 2 Leaf metabolic and structural traits in sink-limited (grey points) and control (black points) plants. A, leaf water content (LWC); B, leaf mass per unit area (LMA); C, free amino acids; D, glucose; E, fructose; F, sucrose; G, starch; H, total non-structural carbohydrates (TNC); I, Protein. TNC is the sum of glucose, fructose, sucrose and starch. For leaf metabolic traits (C,D,E,F,G,H,I), n = 6 plots with the exception of measurements made 9, 15 and 25 days into treatment when 5 sink-limited plots and 4 control plots were measured. The data for each plot represents an average value from three harvested leaves. For leaf structural traits (A,B), sampling was identical to leaf metabolic traits with the exception of the final time point on which samples were taken for LWC and LMA only, with n = 2 sink-limited and n = 2 control plots. Means ± standard error of plot-level data are shown.

Figure 3 An aerial image of the field (A) shows a prolonged green phenotype in plants which underwent the sink removal treatment. Plot positions are indicated as follows: green = control, beige = drought, pink = sink removal (B). Drought plants were not used in the main study but were included in PLSR model building to increase model performance (see Methods).

Figure 4 Photographs show spectral data collection at leaf (A) and canopy (B) levels. The spectroradiometer is referenced using a white Spectralon® disc for leaf-level measurements, and a white Spectralon® plate for canopy-level measurements, shown in (B).

Figure 5 Partial least squares regression (PLSR) models demonstrate that spectral data may be used to predict leaf structural (A,B) and metabolic (C,D,E,F,G,H,I) traits at both leaf and canopy scales. Each plot shows the validation results for leaf-level (black points) and canopy-level (grey points) models. Plots show the relationship between observed traits from traditional measurements, and predicted traits derived from spectral data. The dashed line shows the 1:1 relationship and R^2 and RMSE values are provided for each model (black and grey text is used for leaf and canopy models, respectively). For leaf PLSR models, calibration datasets included between 178 and 192 datapoints, and validation datasets included between 47 and 50 datapoints (Table 2). For canopy PLSR models, there were

between 67 and 71 calibration datapoints, and between 18 and 20 validation datapoints (Table 2).

Figure 6 Linear discriminant analysis (LDA) using measured leaf traits and partial least squares discriminant analysis (PLS-DA) using raw spectral data measured at leaf and canopy scales successfully predict whether or not plants were sink-limited. The grey line represents the 50% (equal to chance) detection rate. AUC-ROC values are reported in the text. Asterisks indicate when a subset of data was used, with aligned leaf and canopy measurement periods, to facilitate comparison with canopy-level data, for which collection began later than leaf-level data.

Figures

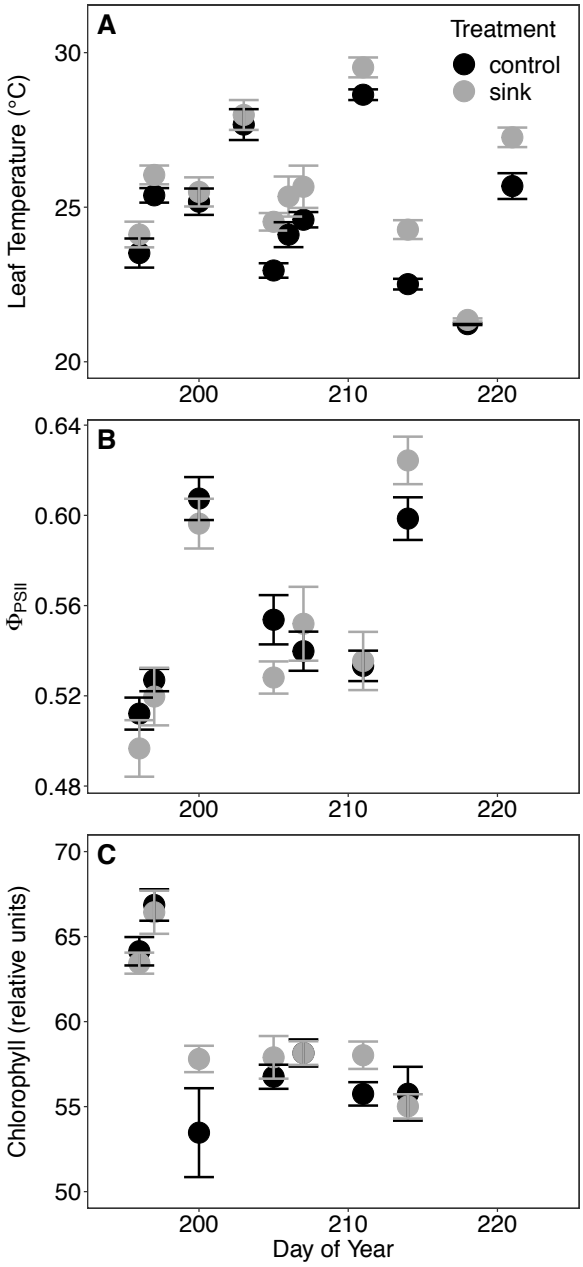


Figure 1 Leaf temperature was higher in sink-limited plants (grey points), but photosynthesis-related traits did not change compared to controls (black points). A, leaf temperature; B, efficiency of photosystem II (ΦPSII); C, relative chlorophyll content. For leaf temperature (A), n = 6 plots each reporting an average value of 3-6 reps per plot measured at leaf and/or canopy scales on any given date; the final two time points show the average of 5 sink-limited plots and 4 control plots. For MultispeQ measurements (B,C), n = 6 plots with an average of 3 reps taken for the plot-level value. Means ± standard error of plot-level data are shown.

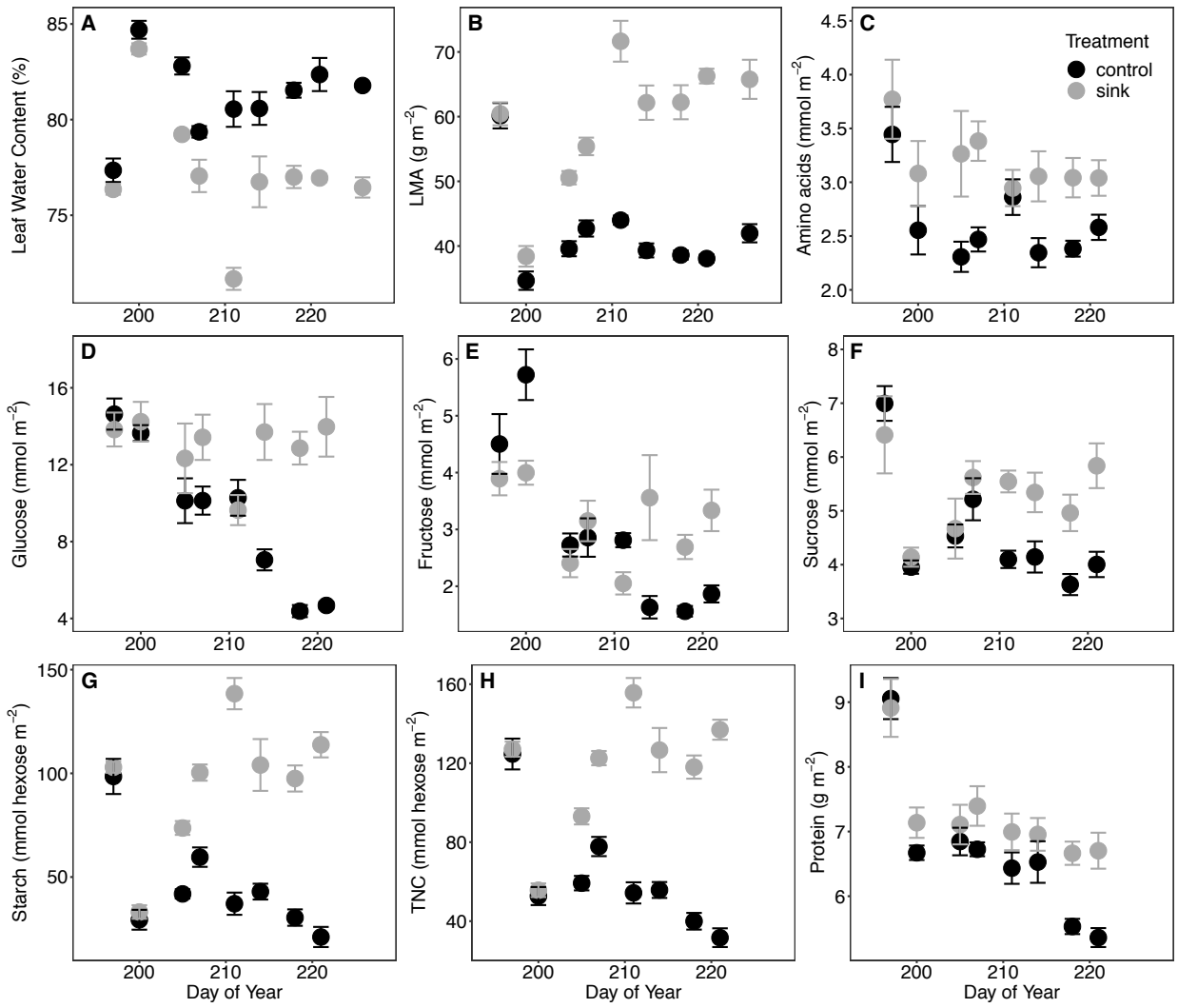


Figure 2 Leaf metabolic and structural traits in sink-limited (grey points) and control (black points) plants. A, leaf water content (LWC); B, leaf mass per unit area (LMA); C, free amino acids; D, glucose; E, fructose; F, sucrose; G, starch; H, total non-structural carbohydrates (TNC); I, Protein. TNC is the sum of glucose, fructose, sucrose and starch. For leaf metabolic traits (C,D,E,F,G,H,I), n = 6 plots with the exception of measurements made 9, 15 and 25 days into treatment when 5 sink-limited plots and 4 control plots were measured. The data for each plot represents an average value from three harvested leaves. For leaf structural traits (A,B), sampling was identical to leaf metabolic traits with the exception of the final time point on which samples were taken for LWC and LMA only, with n = 2 sink-limited and n = 2 control plots. Means ± standard error of plot-level data are shown.

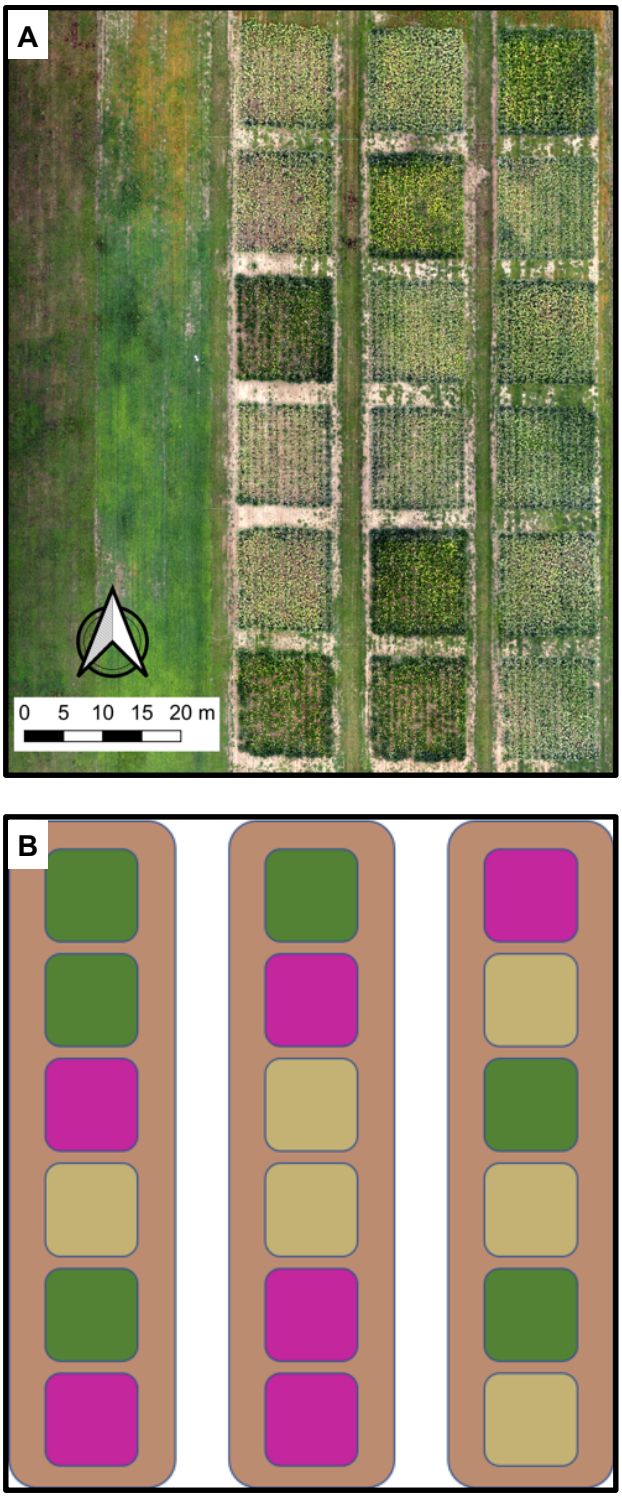


Figure 3 An aerial image of the field (A) shows a prolonged green phenotype in plants which underwent the sink removal treatment. Plot positions are indicated as follows: green = control, beige = drought, pink = sink removal (B). Drought plants were not used in the main study but were included in PLSR model building to increase model performance (see Methods).

Figure 4 Photographs show spectral data collection at leaf (A) and canopy (B) levels. The spectroradiometer is referenced using a white Spectralon® disc for leaf-level measurements, and a white Spectralon® plate for canopy-level measurements, shown in (B).

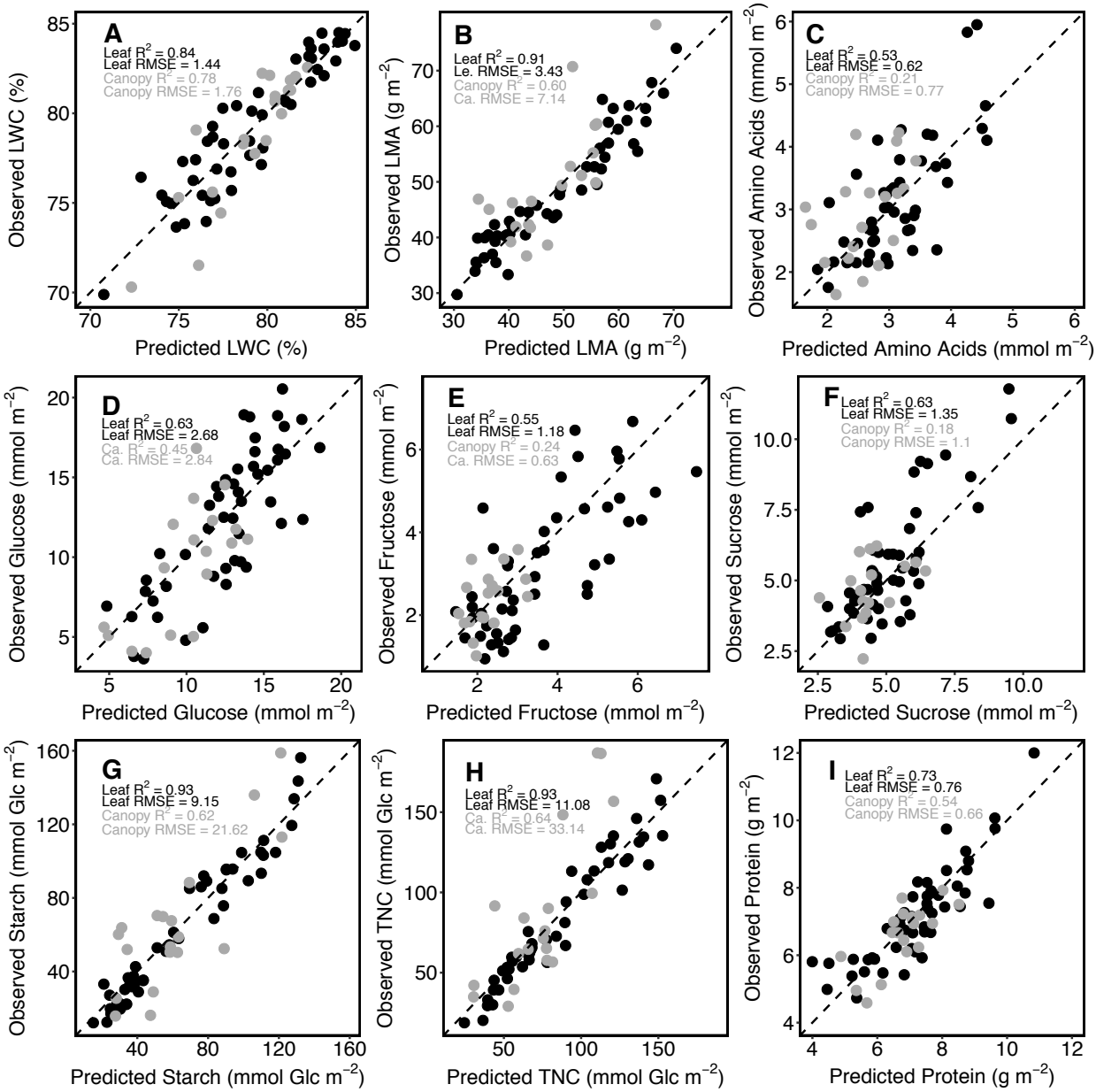


Figure 5 Partial least squares regression (PLSR) models demonstrate that spectral data may be used to predict leaf structural (A,B) and metabolic (C,D,E,F,G,H,I) traits at both leaf and canopy scales. Each plot shows the validation results for leaf-level (black points) and canopy-level (grey points) models. Plots show the relationship between observed traits from traditional measurements, and predicted traits derived from spectral data. The dashed line shows the 1:1 relationship and R^2 and RMSE values are provided for each model (black and grey text is used for leaf and canopy models, respectively). For leaf PLSR models, calibration datasets included between 178 and 192 datapoints, and validation datasets included between 47 and 50 datapoints (Table 2). For canopy PLSR models, there were between 67 and 71 calibration datapoints, and between 18 and 20 validation datapoints (Table 2).

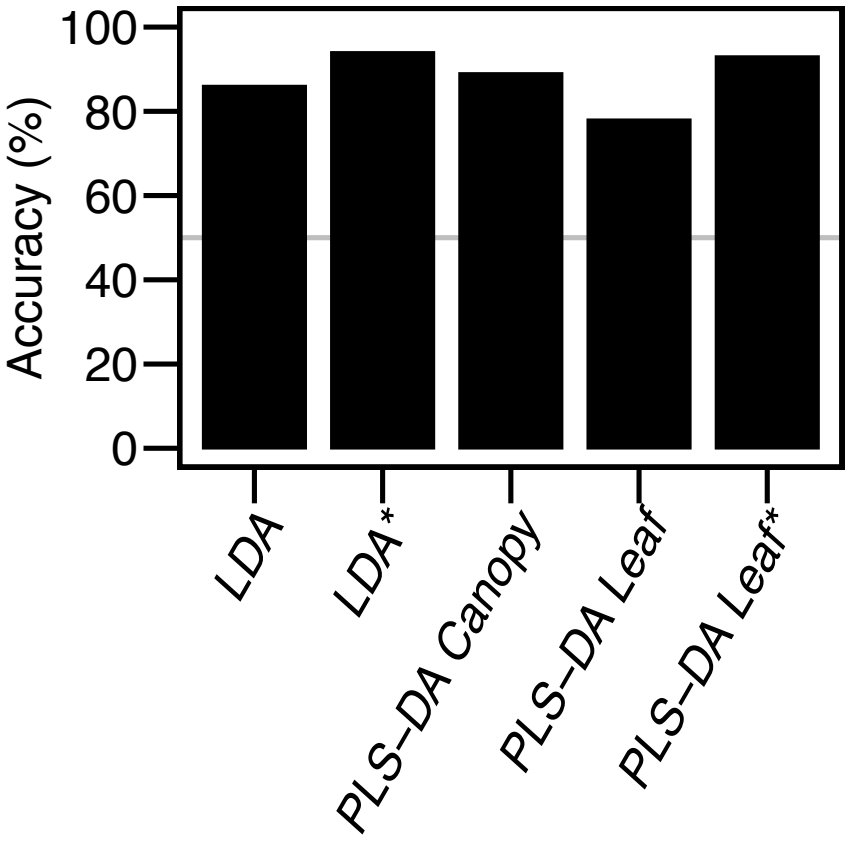
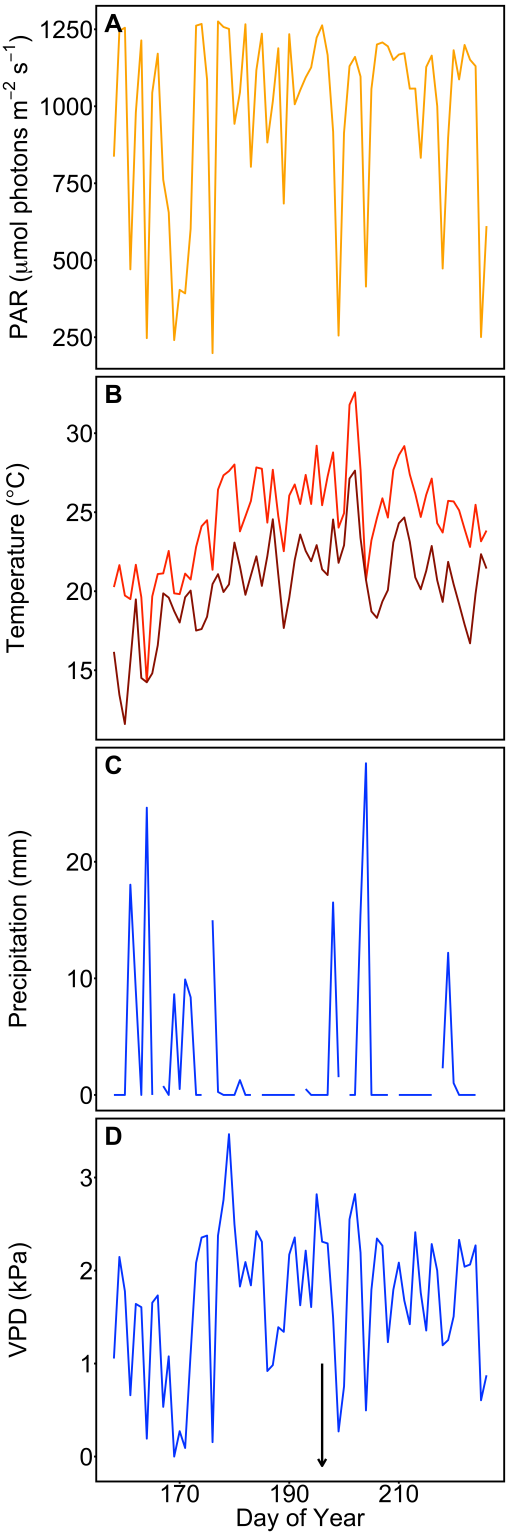
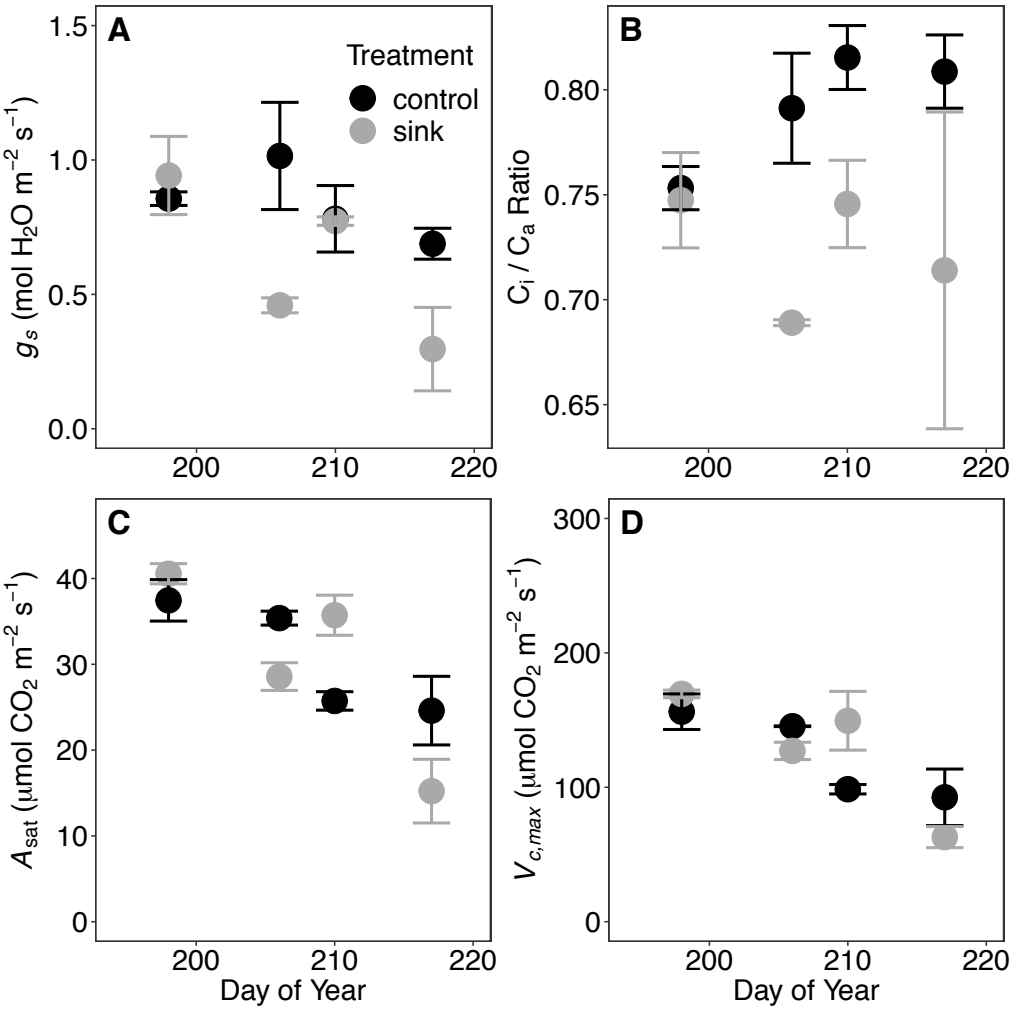


Figure 6 Linear discriminant analysis (LDA) using measured leaf traits and partial least squares discriminant analysis (PLS-DA) using raw spectral data measured at leaf and canopy scales successfully predict whether or not plants were sink-limited. The grey line represents the 50% (equal to chance) detection rate. AUC-ROC values are reported in the text. Asterisks indicate when a subset of data was used, with aligned leaf and canopy measurement periods, to facilitate comparison with canopy-level data, for which collection began later than leaf-level data.

Supplementary Figure 1 Meteorological data for the duration of the experiment from sowing (DOY 158) to final measurement (DOY 226), obtained from the onsite weather station at Brookhaven National Laboratory. A, Mean photosynthetically active radiation (PAR) during daytime hours. B, mean daytime temperature (bright red) and mean nighttime temperature (dark red). C, total precipitation during each 24-hour period. D, maximum daily vapor pressure deficit (VPD) during daytime hours. Initiation of the sink manipulation treatment (DOY 196) is shown by a vertical arrow in panel D.



Supplementary Figure 2 Gas exchange data obtained from A/C_i curves performed on control and sink-limited plants during the course of the experiment. A, stomatal conductance (g_s) at 400 µmol mol^{-1} CO_2 under light-saturated, steady-state conditions; B, ratio of internal and external CO_2 concentrations (C_i/C_a) at 400 µmol mol^{-1} CO_2 under light-saturated, steady-state conditions; C, light-saturated, stable rate of photosynthesis measured at 400 µmol mol^{-1} CO_2 (A_{sat}); D, maximum carboxylation rate of Rubisco ($V_{c,max}$) normalized to 25°C. Plots show means ± standard error. N = 2 plants.



Supplementary Figure 3 Boxplot of Green Chromatic Coordinate (GCC) in control and sink-limited plants at the end of the experiment (DOY 226), obtained from a UAS flight over the experimental field.